REFERENCES & BIBLIOGRAPHY

A Kavitha, S, Suseela, G, G Kapilya (2013),"Big Data", CSI Communication, ISSN 0970-647X, Issue No.1 vol.37, pp. 19.

Albert Yu, Pankaj K. Agarwal, Jun Yang,(2012) "Subscriber Assignment for Wide-Area Content-Based Publish/Subscribe," *IEEE Transactions on Knowledge and Data Engineering*, vol. 24, no. 10, pp. 1833-1847.

Arnab Nandi, Cong Yu, Philip Bohannon, Raghu Ramakrishna(2012),” Data Cube Materialization and Mining over MapReduce”. IEEE Trans. Knowl. Data Eng. 24(10): 1747-1759

Bhandarkar, M. (2010, April).MapReduce programming with apache Hadoop.In*Parallel& Distributed Processing (IPDPS), 2010 IEEE International Symposium on* (pp. 1-1). IEEE

Borthakur, D., Gray, J., Sarma, J. S., Muthukkaruppan, K., Spiegelberg, N., Kuang, H., & Aiyer, A. (2011, June). Apache Hadoop goes real time at Facebook. In *Proceedings of the 2011 ACM SIGMOD International Conference on Management of data* (pp. 1071-1080).ACM.

Das, S., Sismanis, Y., Beyer, K. S., Gemulla, R., Haas, P. J., & McPherson, J. (2010, June). Ricardo: integrating R and Hadoop. In *Proceedings of the 2010 ACM SIGMOD International Conference on Management of data* (pp. 987-998). ACM

Destercke, Sebastien, Buche, Patrice, and Charnomordic, Brigitte(2013), Evaluating Data Reliability: An Evidential Answer with Application to a Web-Enabled Data Warehouse; pp. 92-105.

Dr. Milind Bhandarkar (2013) " Big Data Systems: Past, Present & (possibly) Future " CSI Communication, ISSN 0970-647X, Issue No.1 vol.37, pp. 7-8.

Dr. Patil (2012), Data Jujitsu: The Art of Monetizing Data, ISBN: 9781449341152.

Elena Baralis, Luca Cagliero, Paolo Garza(2013) "EnBay: A Novel Pattern-Based Bayesian Classifier," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 12, pp. 2780-2795

Eltabakh, M. Y., Tian, Y., Özcan, F., Gemulla, R., Krettek, A., & McPherson, J. (2011).CoHadoop: flexible data placement and its exploitation in Hadoop.*Proceedings of the VLDB Endowment*, *4*(9), 575-585.

Francois Goasdoue, M-C Rousset (2013), "Robust Module-Based Data Management," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 3, pp. 648-661

Gautam shroff, Lipika Dey, Puneet Agarwal (2013) , Technical Trends Socio-Business Intelligence Using Big Data ", CSI Communication, ISSN 0970-647X, Issue No.1 vol.37, pp. 11-16.

H. Altay Guvenir, Murat Kurtcephe (2012), "Ranking Instances by Maximizing the Area under ROC Curve," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 10, pp. 2356-2366

Huang, Jiun-Long, and Huang, Chen-Che (2012), A Proxy-Based Approach to Continuous

Husain, M. F., Doshi, P., Khan, L., &Thuraisingham, B. (2009). Storage and retrieval of large rdf graph using hadoop and mapreduce. In *Cloud Computing*(pp. 680-686). Springer Berlin Heidelberg.

Ibrahim, S., Jin, H., Lu, L., Qi, L., Wu, S., & Shi, X. (2009). Evaluating mapreduce on virtual machines: The hadoop case. In *Cloud Computing* (pp. 519-528). Springer Berlin Heidelberg.

KashifJaved, Haroon A. Babri, Mehreen Saeed, (2012) "Feature Selection Based on Class-Dependent Densities for High-Dimensional Binary Data, "*IEEE Transactions on Knowledge and Data Engineering*, vol. 24, no. 3, pp. 465-477.

Leverich, J., & Kozyrakis, C. (2010).On the energy (in) efficiency of hadoop clusters. *ACM SIGOPS Operating Systems Review*, *44*(1), 61-65.

Ling Hu, Wei-Shinn Ku, Spiridon Bakiras, Cyrus Shahabi (2013) "Spatial Query Integrity with Voronoi Neighbors," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 4, pp. 863-876

Liu, X., Han, J., Zhong, Y., Han, C., & He, X. (2009, August).Implementing WebGIS on Hadoop: A case study of improving small file I/O performance on H DFS. In *Cluster Computing and Workshops, 2009.CLUSTER'09. IEEE International Conference on* (pp. 1-8). IEEE.

Location-Based Spatial Queries in Mobile Environments; *Engineering*, vol. 25, no. 2, pp. *260-273.*

Luca Cagliero, Alessandro Fiori(2013)Discovering generalized association rules from Twitter" *IOS Press*, vol. 17, no.4/2013, pp.627-648.

Lushan Han, Tim Finin, Paul McNamee, Anupam Joshi, Yelena Yesha (2013), "Improving Word Similarity by Augmenting PMI with Estimates of Word Polysemy," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 6, pp. 1307-1322.

M. Sivasankari (2014), "A Novel Event Rule Derivation for Processing Uncertain Data Events Using Genetic Network Programming", *International Journal of Computer& organization Trends (IJCOT)*,V9(1):8-12

Manish Bhide, Krithi Ramamritham,"Category-Based Infidelity Bounded Queries over Unstructured Data Streams," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 11, pp. 2448-2462.

Maria Kontaki, Apostolos N. Papadopoulos, Yannis Manolopoulos(2012), "Continuous Top-k Dominating Queries" *IEEE Transactions on Knowledge and Data Engineering*, vol. 24, no.5, pp.840-853.

Mohammad M. Masud, Qing Chen, Latifur Khan, Charu C. Aggarwal, Jing Gao, Jiawei Han, Ashok Srivastava, Nikunj C. Oza(2013), "Classification and Adaptive Novel Class Detection of Feature-Evolving Data Streams," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 7, pp. 1484-1497.

Nandi, C. Yu, P. Bohannon, and R. Ramakrishnan (2011), "Distributed cube materialization on holistic measures", in Proc. ICDE, pp.183-194.

Pavel Shvaiko, Jérôme Euzenat (2013)"Ontology Matching: State of the Art and Future Challenges," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 1, pp.158-176

Salvador Garcia, J. Luengo, José Antonio Sáez, Victoria López, F. Herrera (2013), "A Survey of Discretization Techniques: Taxonomy and Empirical Analysis in Supervised Learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 4, pp. 734-750

Segev Wasserkrug, Avigdor Gal, OpherEtzion, Yulia Turchin (2012), "Efficient Processing of Uncertain Events in Rule-Based Systems," *IEEE Transactions on Knowledge and Data Engineering*, vol. 24, no. 1, pp. 45-58.

Shafer, J., Rixner, S., & Cox, A. L. (2010, March). The Hadoop distributed file system: Balancing portability and performance. In *Performance Analysis of Systems & Software (ISPASS), 2010 IEEE International Symposium on* (pp. 122-133). IEEE.

Shailesh Kumar Shivakumar (2013) " Big Data-A Big game changer " CSI Communication, ISSN 0970-647X, Issue No.1 vol.37, pp. 9-10.

Shvachko, K., Kuang, H., Radia, S., &Chansler, R. (2010, May). The hadoop distributed file system. In *Mass Storage Systems and Technologies (MSST), 2010 IEEE 26th Symposium on* (pp. 1-10). IEEE.

Tao Chen, Lei Chen, M.TamerOzsu, Nong Xiao(2013), "Optimizing Multi-Top-k Queries over Uncertain Data Streams," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 8, pp. 1814-1829

Taylor, R. C. (2010). An overview of the Hadoop/MapReduce/HBase framework and its current applications in bioinformatics. *BMC bioinformatics*, *11*(Suppl 12), S1.

Thusoo, A., Sarma, J. S., Jain, N., Shao, Z., Chakka, P., Zhang, N., ...& Murthy, R. (2010). Hive-a petabyte scale data warehouse using hadoop. In *Data Engineering (ICDE), 2010 IEEE 26th International Conference on* (pp. 996-1005). IEEE.

Viktor Mayer-Schonberger, Kenneth Cukier (2013), Big Data, Murray Publication, ISBN 978-1-84854-793-3.

Wang, J., Crawl, D., & Altintas, I. (2009, November).Kepler+ Hadoop: a general architecture facilitating data-intensive applications in scientific workflow systems. In *Proceedings of the 4th Workshop on Workflows in Support of Large-Scale Science* (p. 12).ACM.

Yiyao Lu, Hai He, Hongkun Zhao, WeiyiMeng, Clement Yu (2013), "Annotating Search Results from Web Databases," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, no. 3, pp. 514-527.

Zhao, Z.D., & Shang, M. S. (2010, January).”User-based collaborative-filtering recommendation algorithms on hadoop. In *Knowledge Discovery and Data Mining, 2010.WKDD'10. Third International Conference on* (pp. 478-481).

Bibliography of periodicals

1. **CSI** Communication: A monthly by Computer Society of India
2. **Developer IQ** is magazine on appreciating and research on software Development.
3. **OpenSource** as the name suggested it is the complete magazine on open source.

Bibliography of Books

1. Viktor Mayer-Schonberger and Kenneth Cukier(2013), Big Data, ISBN 978-1-84854-793